UNIVERSITÀ DI PISA
DIPARTIMENTO DI FARMACIA

Corso di Laurea Magistrale in
Chimica e Tecnologia Farmaceutiche

TESI DI LAUREA:

Stereoselective synthesis of a new carba sulfonamide pseudomannobioside as possible DC-SIGN ligand.

Relatore: Dott.ssa Valeria Di Bussolo

Candidata: Serena Fortunato

ANNO ACCADEMICO 2014-2015

Index

ABSTRACT

Dendritic cells (DC) capture microorganisms that enter peripheral mucosal tissues and then migrate to secondary lymphoid organs, where they present these in antigenic form to resting T cells and thus initiate adaptive immune responses. A DC-specific C-type lectin, DC-SIGN, is highly expressed on DC present in mucosal tissues and binds to the HIV-1 envelope glycoprotein gp120. DC-SIGN efficiently captures HIV-1 in the periphery through a terminal carbohydrate recognition domain (CRD).[1] The main carbohydrate ligand recognized by DC-SIGN is the high mannose glycan $Man_9(GlcNAC)_2$, a branched oligosaccharide presented in multiple copies by several pathogen glycoproteins. In the branched oligosaccharide, the terminal disaccharide portion Manα1-2Man binds DC-SIGN almost as efficiently as the entire high mannose glycan $Man_9(GlcNAc)_2$. This aspect suggests an important role of nonreducing end Manα1-2Man fragment of Man_9 in DC-SIGN recognition[2].

Recently, it has been settled on a new class of DC-SIGN antagonists[2]: pseudodisaccharides in which the reducing mannose unit is replaced by a conformationally restricted dimethyl cycloexandicarboxylate.

The main purpose of this thesis project is to operate a stereoselective synthesis of new DC-SIGN antagonist, characterized by the remarkable presence of a real D-carbamannose unit with a hydrophobic portion, represented by a *p*-toluensulfonamide group at C(4) position. To build the pseudomannobioside, a new carbamannose unit, as glycosyl acceptor was connected to an appropriate glycosyl donor, represented by a manno-trichloroacetimidate (TCA) and the resulting pseudodisaccharide was subjected to different elaboration in order to obtain the desired pseudodisaccharides fully-*O*-deprotected, with an ethoxy-amino chain on pseudo-anomeric position.

INTRODUCTION

1. VIRAL PROTEINS AND SURFACE CARBOHYDRATES, TARGET OF DC-SIGN

1.1 Structures of Viral Surface

Enveloped viral particles usually contain different types of surface proteins the majority of which is modified by the addition of *N*-linked, *O*-linked carbohydrate chains or both.

Typically, type I membrane proteins are characterized by a cleaved *N*-terminal signal peptide, a *C*-terminal hydrophobic anchor sequence and a small *C*-terminal cytoplasmic domain. Type 2 proteins have an uncleaved *N*-terminal signal-anchor sequence and type 3 proteins with multiple membrane spanning domains also exist. However, the major proportion of such glycoproteins is usually situated on the external side of viral envelope.

Totally viral envelope has shown to be extensively modified by *N*-linked glycosylation; however, the presence of *O*-linked carbohydrates on the glycoprotein has not been firmly established. *N*-terminal attachment site is a constant region denoted principally by an Asn residue. Instead, Ser and Thr residues situated either in the proline-rich region of the molecule or in close proximity to the previous Asn, are shown to be *O*-glycosylated.[3]

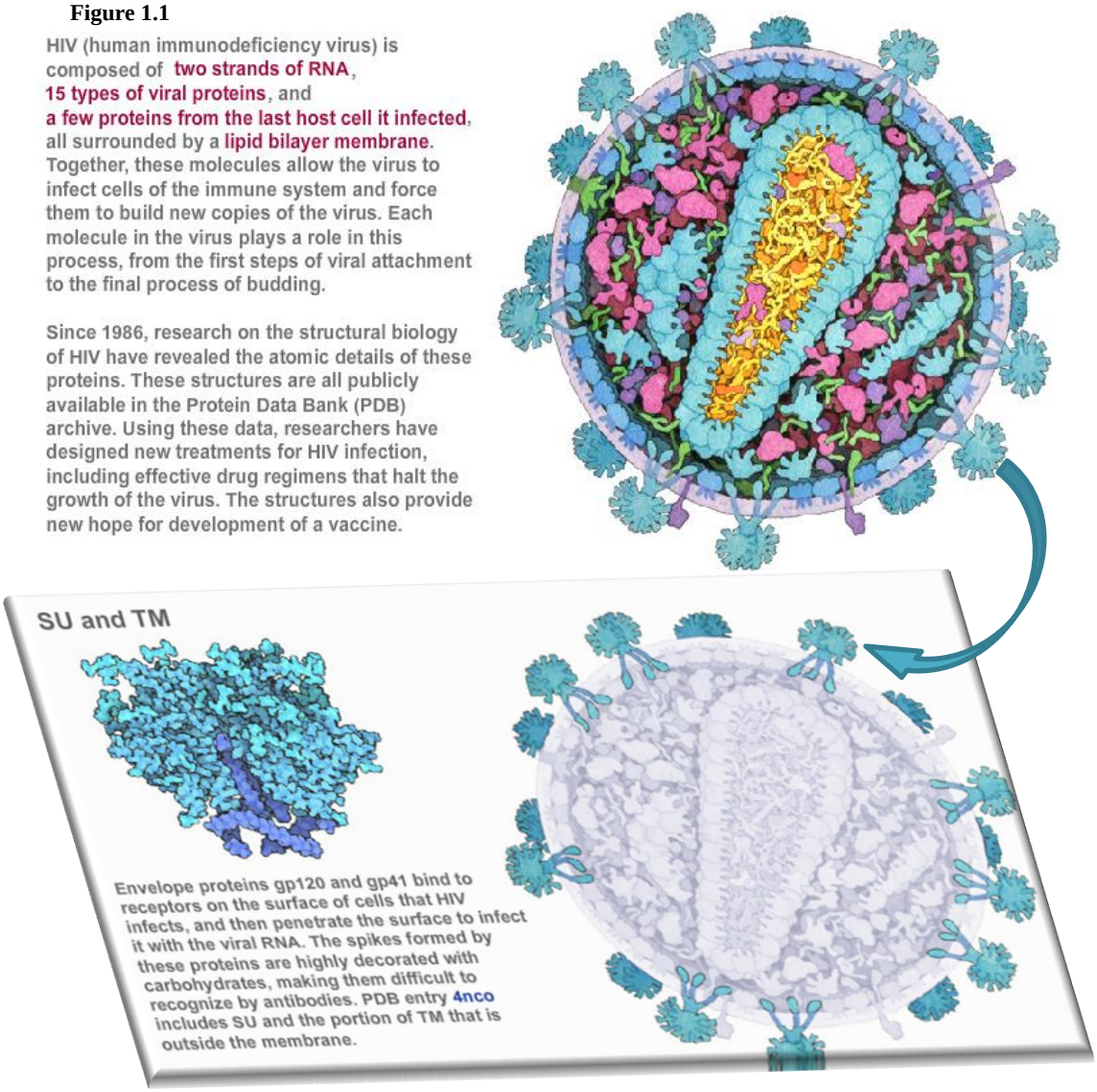
Knowledge of carbohydrate moieties linked to virus surface glycoproteins comes from a detailed structural analysis of some viruses having great biological and biomedical importance, such as Murine leukemia viruses, Marburg virus (MBGV), Hepatitis virus and finally as HIV. It demonstrates that specific *N*-glycosylation sites link above all high-mannose, hybrid or complex type di-, tri- and tetra-antennary *N*-glycans. The other carbohydrate constituents are represented by fucose and/or galactose residues (Table 1).[4] To offer an idea of surface cell glycoproteins concentration, I take care above all of many studies covered on the human immunodeficiency virus type 1 (HIV-1) envelope glycoproteins.

1.1.1 The HIV-1 envelope glycoprotein

Figure 1.1

HIV (human immunodeficiency virus) is composed of **two strands of RNA**, **15 types of viral proteins**, and **a few proteins from the last host cell it infected**, all surrounded by a **lipid bilayer membrane**. Together, these molecules allow the virus to infect cells of the immune system and force them to build new copies of the virus. Each molecule in the virus plays a role in this process, from the first steps of viral attachment to the final process of budding.

Since 1986, research on the structural biology of HIV have revealed the atomic details of these proteins. These structures are all publicly available in the Protein Data Bank (PDB) archive. Using these data, researchers have designed new treatments for HIV infection, including effective drug regimens that halt the growth of the virus. The structures also provide new hope for development of a vaccine.



The HIV Env glycoprotein precursor, gp160, is a highly glycosylated protein of approximately 850 amino acids. During intracellular transport, the gp160 polyprotein is cleaved into two subunits that remain associated: gp41, which contains ecto-, transmembrane, and cytoplasmic domains, and gp120, which is non-covalently linked to the ectodomain of gp41 (Figure 1.1).

The number of potential *N*-glycosylation sites on gp120 is on the order of 20, but the oligosaccharide structures are far more numerous. In fact, gp120 is extensively modified by *N*-linked glycosylation as much as 50% of the molecular mass of the major envelope glycoprotein by carbohydrates in the form of oligosaccharide chains attached to the polypeptide backbone. Thus, it represents one of the most heavily glycosylated molecules in nature.

1.1.2 Oligosaccharides on N-glycosylation sites

Protein *N*-glycosylation is an important co-translational modification process wherein short sugar chains are covalently attached to the amide group of asparagine (*N*) residue in the amino acid chain. Although asparagine occurs frequently in the protein chain, *N*-glycosylation requires asparagine to be present in special motifs, in NXS/T sequons, the potential sites of *N*-glycosylation where X is any amino acid except proline which is avoided due to conformational hindrance and the third residue is either serine or threonine. Furthermore, *N*-glycosylation occurs only on some sequons found in membrane-bound or secretory proteins which are exposed to the enzyme oligosaccharyl-transferase in the lumen of endoplasmic reticulum. To know the number and distribution of *N*-glycosylated sequons, gp120 sequence is divided in two (*N* and *C* terminal) parts. Two rational methods are used: first, the gp120 has two parts namely the inner and outer domains. The bulk of the *N* terminal part of the sequence forms the inner domain and the *C* terminal part forms the outer domain. The amino acid chain makes a transition from inner to outer domain. Second, it is known that the two domains are unequally glycosylated with a higher proportion occurring in the outer domain. It may be thought that the outer domain is much more exposed to solvent and has a number of regions, which interact with host proteins as compared to the inner domain, much of it is facing gp41 or inner domains of gp120 monomers within the trimeric complex. Each monomer in non-covalent association with gp41, the transmembrane viral envelope glycoprotein anchor, interacts with its primary receptor CD4 glycoprotein of the host T-lymphocyte.

Some study found no significant change over time in the number of NXS/T sequons in the whole gp120 molecule, which contains an average of 26 NXS/T sequons. The gain or loss of sequons may influence the gp120/virus. For example, the glycan structures are quite bulky (~2000 Da) and therefore, the additional sequons may not provide a selective advantage due to conformational instability or functional redundancy. Similarly, the loss of sequons may have fitness costs, as they lead to failure of infectivity/immune evasion. In addition, in more variable regions of the envelope, amino acid exchanges may cause the addition or deletion of glycosylation site. It is hypothesized that simply under selection pressures, the gp120 must naturally undergo rapid changes with respect to its *N*-glycosylation sequons. For instance, under selection pressure by mannose-specific lectins, resistant HIV-1 strains showed marked depletion (up to eight out of 22) of the *N*-glycosylation sequons in gp120 regions distant from interacting sites. Fluctuations in the variable regions of the gp120 itself has been shown to change the *N*-glycosylation sequons in the early stages of viral infection. Likewise, it is reasonable to expect a long-term trend in the number of *N*-glycosylation sequons in gp120 under the selection pressure from the host immune system or anti-retroviral drugs since the passage of HIV to human hosts.[5]

Finally, other studies show a very clear directional change in the density (or number) of sequons in the two domains of gp120. The NXT sequon density is decreasing in the outer domain and the NXS sequon density is increasing. The increase in NXS density is significant even when gp120 molecule is considered as a whole. But it is interesting to note that there is no net change in the NXT sequon density in the whole gp120 molecule. This is because the decrease in NXT density in outer domain is compensated by significant increase in NXT density in inner domain over the time period between 1981 and 2009. It may be recalled that previous studies failed to notice these intricate trends, instead reported mere fluctuations in the sequon numbers in gp120 over time.

Leaving out the changes regarding NXS/T sequons, some studies were addressed on recombinant gp120 produced in chronically infected T lymphocytes (H9 cells) and transfected Chinese hamster ovary (CHO) cells. A different structural characterization of carbohydrates sequences on gp120 was identified including high mannose-type (Man_5 to Man_9 structures amounting to -33%) and hybrid-

type (4%) chains as well as four categories of complex-type chains (mono-, bi-, tri-, and tetraantennary) with or without *N*-acetyllactosamine repeats and with or without core region fucose residues among which digalactosyl biantennary structures predominated (34%). Altogether, 29 structures were identified after desialylation.

The process of *N*-linked glycosylation can result in diverse structures that may be divided into the three categories: oligomannose, hybrid, and complex. Each category shares a common $Man_3GlcNAc_2$ pentasaccharide stem (where Man is mannose and GlcNAc is *N*-acetylglucosamine), to which up to six mannose residues are attached in oligomannose *N*-glycans, while complex *N*-glycans are usually larger and may bear various sizes and numbers of antennae. Glycan synthesis begins in the endoplasmic reticulum, where *N*-linked oligomannose precursors $Glc_3Man_9(GlcNAc)_2$ (Glc is glucose) are transferred cotranslationally to the free amide of the asparagine in a sequon Asn-X-Thr/Ser.

The actual number of oligosaccharides is much greater: there was evidence that among the hybrid- and complex-type chains, all but 6% contained sialic acid at C-3 in terminal galactose residues, and partially sialylated forms of the bi- and multiantennary chains were present.

The marked variability of the glycosylation details is considerably influenced by the host cell.[6] The oligosaccharide profiles were compared for the envelope glycoproteins of different virus isolates that differ in their biological properties and amino acid sequences of their *env* genes. When the virus isolates were propagated in the same host cells, their *N*-glycan patterns were similar, whereas there were marked differences when propagated in different host cells. For example, the HIV-1 gp120 produced in insect cells was found to have a repertoire of high mannose type but no complex type *N*-glycans. Instead, gp160, and so gp120, produced in the human lymphoblastoid cell line Molt-4 demonstrated the presence of sulfate 6-linked to the *N*-acetylglucosamine of complex type chains.

1.1.3 Oligosaccharides on O-glycosilation sites

There were several investigations of the presence of *O*-linked oligosaccharides on the HIV-1 envelope glycoprotein, but it has not been conclusively established. *O*-linked glycosylation has been demonstrated, however, for other retroviral glycoproteins, including Friend spleen focus-forming virus (SFFV), Friend mink cell focus-forming virus (FrMCF), Rauscher murine leukemia virus (R-MuLV), and feline leukemia virus. The *O*-linked carbohydrates were localized above all to the surface subunits of the FrMCF glycoproteins.

O-linked glycosylation is a commonly occurring form of post-translational modification; so it begins with the addition of the core sugar GalNAc in an α-glycosidic linkage to a serine or threonine residue. The basic structure may then be enhanced by the subsequent addition of galactose, *N*-acetylglucosamine, fucose, and sialic acid by glycosyltransferases found in the Golgi apparatus. The structures of *O*-linked carbohydrates are varied, and there is no common carbohydrate core or processing pathway similar to that involved in *N*-linked glycosylation.

Indirect evidence supporting the presence of *O*-linked oligosaccharides on the HIV-1 envelope glycoprotein has previously been obtained: it was observed a difference in the pI between envelope glycoprotein that had been deglycosylated with PNGase F and envelope glycoprotein that had been treated with neuraminidase prior to deglycosylation with PNGase F, suggesting that sialated *O*-linked carbohydrates may be present on the HIV-1 envelope glycoprotein. Our observation that treatment with neuraminidase followed by *O*-glycosidase causes a decrease in the apparent molecular weight of gpl20 and gpl60 correlates with their data and suggests that both the precursor and surface unit forms of the glycoprotein contain *O*-linked glycans.

Likewise, immuno-precipitation experiments carried out with anticarbohydrate monoclonal antibodies demonstrate *O*-linked carbohydrates presence. These antibodies were found to immunoprecipitate gpl20 and inhibit infection and syncytium formation in $CD4^+$ lymphocytic cell lines. However, the common *O*-linked sequence Galβ1-3GalNAc, with or without terminal sialic acid, could not be detected on natural gp120 derived from infected H9 cells, or on recombinant gp 120 from CHO cells, a cell line which is reversibly defective in *O*-linked glycosylation because of its inability to synthesize UDP-Gal and UDP-GalNAc. As described before, glycosylation is host cell specific and is dependent on the presence of glycosyltransferases, which vary among different

cell types. In addition, variation in glycosylation may even be seen within a single cell line, as observed by Mizuochi et al.,[6] who analyzed the *N*-linked carbohydrate structures present on the envelope glycoprotein of HIV-IIIB isolated from H9 cells.

1.2 Functions of Viral Surface Glycoproteins

Most viral envelope glycoproteins are essential during the viral life cycle. They permit virion's assembly and defend it from enzymatic attack; but their key role in human is to induce a humoral immune response.

First of all, it is necessary explain that target cells for viral glycoproteins are various: for example, HIV can infect CD4 T cell, macrophages and microglial cells. This is possible because viral glycoproteins chains are characterized by many moieties, each of which interacts with a specific cell receptor. Our attention was focused on dendritic cell-specific ICAM-3 grabbing nonintegrin (DC-SIGN) receptor that interacts with many viruses, as HIV, SIV, Ebola virus, Hepatitis C, Dengue, SARS, and also some bacteria (Mycobacterium tuberculosis, Helicobacter pylori). So, mediating the first contact with host cell receptor, virus binding and subsequent fusion of viral and cell membranes, glycoproteins direct starting infection.

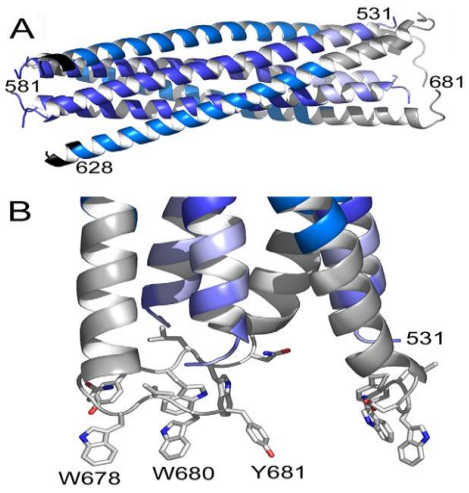


Figure 1.2
A) Ribbon representation of gp41;
B The exposure of aromatic side chains Trp 678, Trp 680 and Tyr 681 towards the membrane.

Thus, virion can penetrate into the host cell, where makes its genome accessible for transcription. For some viruses, a single membrane glycoprotein is

responsible for both functions, whereas for others, receptor binding and fusion activities reside in different glycoproteins. For examples, Human Immunodeficiency Virus (HIV) envelope glycoproteins, gp120 and gp41 are both vital for virus entry into host cell: gp120 is essential for viral adhesion to specific surface receptors of various cell types, although gp41 interacts with the specific cell membrane in a second time. After conformational and position change of the first glycoprotein, gp41 (showed in Figure 1.2) allows the beginning of membranes fusion process and virion entry into host cell cytoplasm.

However, attachment of an envelope glycoprotein to a cell surface requires a specific binding among viral polypeptide chain and host cellular receptor. In fact, the HIV-1 pathogenesis, to a large extent, has been attributed to the structural plasticity of the gp120, and in particular, to the variability of *N*-glycosylation, that allows specific interactions.[7]

Since 1992, other studies have been focused on the influence of *N*-linked oligosaccharides on the transmembrane glycoprotein gp41of HIV-1 on viral infectivity. In the first investigations, some amino acid changes caused a severe impairment of viral infectivity toward a CD4-positive cell line; in the second of these, variable effects on infectivity toward different CD4-positive cells were observed particularly with respect to fusion activity. More recent investigations have reached the conclusion that the glycosylation site at position 621 is the most important for fusion activity, and that gp41 lacking *N*-glycans becomes arrested in the Golgi apparatus. From the observed results, lack of saccharides or critical change in the protein folding and conformation could direct the impairment of binding affinity.[8]

1.2.1 Functions of Carbohydrate Substituents in N-glycosylation sites

From recent studies, *N*-glycosylation is the prerequisite of a number of proteins properties such as solubility, stability and turnover, secretion, protease resistance, protein-protein interaction/recognition and immunogenicity, and hence has an immense biological importance.

The key role of these carbohydrates in protein function and immune recognition is that to mask the antigenicity of the polypeptide backbone; at the same time they guide the presentation of the whole glycoprotein to antigen- presenting cells.

a. Glycosylation can influence the protein active form

In order to determine which specific *N*-linked glycans are critical for envelope protein function or immune escape, several recent studies on HIV envelope glycoprotein have been directed from the *Departments of Pathology, Cell Biology, and Immunobiology, Yale University School of Medicine* of Connecticut on individual or multiple mutations of glycosylation site (as described above).[7a] As result of peptide portion changes, the presence of a variety of carbohydrates is especially critical during early steps of envelope protein folding and cleavage, but once envelope achieves its final conformation, glycosylation is less critical. So, oligosaccharides at highly conserved *N*-glycosylation sites appear to be important for the glycoprotein biologically active form. It is the prerequisite for suitable binding to the host cell receptors.

b. Glycosylation can influence antigenicity and immunogenicity

Carbohydrate-mediated reactivity of gp120 with two proteins of the host was documented thus far. The first is with the serum lectin known as mannose-binding protein, of which will be discussed in greater detail below. The second carbohydrate-mediated interaction, shown with gp120, is with the endocytosis receptor of human macrophage membranes. High affinity binding may lead to viral uptake by macrophages irrespective of the presence of the membrane-associated CD4 receptor.

Some examples of antigenic determinants of the viral envelope include the accessible mannose or fucose residues, the *N*-acetylglucosamine or linear and branched poly-*N*-acetyllactosamine backbone sequences.

Specific envelope glycosylation sites also appear to have an important role in modulating the antibody response.[9] It is the outer domain, divided into a neutralizing face and an immunologically silent face: oligomannose glycans cluster tightly on the silent face of gp120, while complex glycans

flank the gp120 receptor binding sites of the neutralizing face, forming a protective "fence" against Nabs.

For instance, a replace of *N*-linked glycan in the HIV-1_{BRU} envelope V1 region can make the virus more resistant to neutralization by anti-V3 antibodies. Instead, HIV IIIB *env* clones lacking an *N*-glycan in the V3 loop of envelope protein exhibited increased sensitivities to V3 loop-specific monoclonal antibodies (MAbs) and soluble CD4. By masking an immunodominant epitope in the V3 loop with additional *N*-linked carbohydrates, the antibody response can be shifted from the V3 epitope to the V1 epitope.

One of the most dramatic effects of carbohydrate removal from an envelope glycoprotein was reported from many studies with simian immunodeficiency virus (SIV): these results indicated significant differences between SIV and HIV with regard to the roles of glycans. Most mutants, lacking of a number of *N*-linked glycosylation, induced a marked increase antibody and a marked increased neutralizing activity. These results, observed for example in infected Rhesus monkeys[10], demonstrated a role for *N*-linked glycosylation in limiting the neutralizing antibody response to SIV and in shielding the virus from immune recognition.

So, carbohydrates provide a layer of protection against NAb attack. As glycans are considered self, antibody responses against them are thought to be regulated by tolerance mechanisms.

The relatively sparse clustering of complex glycans that form this fence may reflect a trade-off between protecting the underlying functional domains from NAbs thanks to the high density while at the same time permitting sufficient flexibility for the refolding events associated with receptor binding and fusion. Conversely, the dense clustering of oligomannose glycans on the silent domain may be important for ensuring immune protection and/or in creating binding sites for lectins such as DC-SIGN.

1.2.2 Functions of Carbohydrate Substituents in O-glycosylation sites

The function of *O*-linked glycosylation is not well understood. Some studies proposed that the *O*-glycosylation sites may play a role in the stability of cell surface glycoproteins, and prevent their proteolysis.

A recent report by Overbaugh and Rudensky demonstrates that simian immunodeficiency virus variants isolated late in the progression to simian AIDS contain envelope sequences rich in serine and threonine, and they suggest that these changes may be accompanied by increased glycosylation, including potential sites for the addition of *O*-linked glycans. Such addition of oligosaccharides may play a role in escape of viruses from immune surveillance mechanisms by masking antigenic epitopes on the glycoprotein. But a specific role for these structures is yet to define.

1.3 *Immune response and interaction with target cells*

The human immunodeficiency viruses (HIV-1 and HIV-2) and simian immunodeficiency viruses (SIVs) are the etiologic agents of acquired immunodeficiency syndrome (AIDS) in their respective human and simian hosts.[10a]

Typically, infection with primate immunodeficiency viruses is characterized by an initial phase of high-level viremia, followed by a long period of persistent virus replication at a lower level. Viral persistence occurs despite specific antiviral immune responses, which include the generation of neutralizing antibodies.

The primate immunodeficiency viruses, like all retroviruses, are surrounded by an envelope consisting of a host cell lipid bilayer and virus-encoded envelope glycoproteins. Fusion process between these two membranes is very important to the viral evolution in the host. All the viral components are synthetized in the infected cell from genes and retroviral enzymes, coming from viral cells, and then are modified in the Golgi apparatus: gp160 is cleaved by a cellular protease

generating the mature envelope glycoproteins (gp120 and gp41), carbohydrates on the gp120 glycoprotein are modified by the addition of complex sugar and then the assembly virion exit from the host cell.

The exposed location of the proteins on the virus allows them to carry out their function but also renders them uniquely accessible to neutralizing antibodies. Thus, virus replication and immune response drive the evolution of the envelope glycoproteins and continue to do so within each infected host.

Many cell surface proteins, including adhesion molecules, are incorporated into HIV-1 virions along with the envelope glycoprotein complexes. These host cell-derived molecules can assist the attachment of viruses to potential target cells. Virus attachment also involves the interaction of the gp120 envelope glycoproteins with specific receptors, as the CD4 glycoprotein and members of the chemokine receptor family. The CD4 glycoprotein is expressed on the surface of T lymphocytes, monocytes, dendritic cells, and microglia, the main target cells for primate immunodeficiency viruses in vivo. The adhesion to CD4 receptor is not the only one: among the viral attachment factors there is a C-type lectin, DC-specific intercellular adhesion molecule 3 (ICAM-3)-grabbing nonintegrin (DC-SIGN, also known as CD209), which mediates viral transmission with high affinity to the envelope glycans. Attachment to DC-SIGN thus enhances viral spread in mucosal dendritic cells and, by taking advantage of the inherent capacity of DCs to migrate into lymphoid tissue, might promote viral dissemination within the host. DC-SIGN and a related molecule, DC-SIGNR, also enhance infection by Ebola virus. DC-SIGN expression is restricted to certain dendritic cells subgroups, dermal and stromal DCs, and it is not expressed on follicular DCs or on skin-resident LCs. Other cell types that express DC-SIGN include monocyte-derived DCs (MDDCs), macrophages and activated B lymphocytes.

Several weeks after virus infection, usually after the initial high level of viremia has decreased, neutralizing antibodies can be detected infected animals or humans. These antibodies neutralize the infecting virus, but often exhibit little or no activity against other strains of virus. A subset of these strain-restricted antibodies recognizes the HIV-1 V3 loop. These antibodies can block also chemokine receptor binding.

2. DC-SIGN AS TARGET RECEPTOR

2.1 Structure and Function of DC-SIGN

DC-SIGN belongs to calcium (Ca^{2+})-dependent lectin group, known as C-type lectins. It is a type II membrane protein with a single carbohydrate-recognition domain (CRD) represented by a C-terminal portion. Indeed, it is characterized by other three domains: the *N*-terminal cytoplasmic tail and transmembrane domain followed by the extracellular part composed of a neck region.[11]

Lectins class that binds to exogenous sugars, includes the mannose-binding proteins (MBPs) and endocytic receptors, such as the macrophage mannose receptor that mediates the innate immune response against invading microorganisms by interacting with cell surface sugar structures. So, DC-SIGN shares specificity for mannose and are involved in pathogen recognition.[12]

Immature dendritic cells in mucosal tissue use DC-SIGN to recognize high-mannose glycans present on the viral envelope glycoprotein, as on gp120 of HIV. This recognition event appears to contribute to infection by promoting viral transmission in a manner dependent on the composition of gp120 glycans. So, DC-SIGN exists only at low levels on blood monocytes and when they leave blood to peripheral tissues the expression of this lectin is extremely increased. Finally, following DC maturation and migration to secondary lymphoid tissues, DC-SIGN levels are down-regulated. This detail of DC-SIGN is important to understand that the lectin doesn't work as a receptor for viral entry into DC; instead binding of the viral envelope glycoprotein to DC-SIGN may induce a conformational change that enables a more efficient interaction with CD4 and/or chemokine receptor (CCR5) promoting resourceful infection.

2.1.1 Structure of CRD and Calcium-mediated interaction

The Ca^{2+}-dependent CRD (Figure 1.3) of DC-SIGN is a long-form polypeptide sustained by four disulphide bridges. It contains two occupied calcium-binding sites, one of which is essential for the polypeptide tertiary structure and the other to coordinate ligand binding, as mannose group sugars.

Usually, the cation links amino acids and oxygen atoms from carbohydrates hydroxyl groups. Indeed, it can play a significant role in the stereospecific recognition of carbohydrates via the relative stereochemistry (*cis* or *trans*) of two adjacent hydroxyl groups. For instance hydroxyl groups corresponding to the 3- and 4-OHs of mannose serve as coordination ligands for the Ca^{2+}. Additional coordination ligands are provided by asparagine and glutamic acid residues that also form hydrogen bonds with the 3- and 4-OH groups of the sugar and with C6-position.

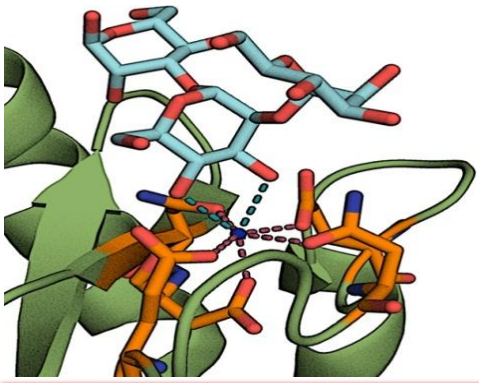


Figure 1.3 *Calcium-mediated interaction of a high mannose N-glycan (cyan) with the human DC-SIGN carbohydrate recognition domain (green/orange)*

Moreover, DC-SIGN binding site shows Val^{351} residue as the main amino acid residue that participates in carbohydrate binding through van der Waals interactions. In addition to the ligand interaction with the primary binding site centered around the Ca^{2+} ion, the ligand forms additional contacts with the surface of DC-SIGN.[13]

Recent studies as Cross-linking, Equilibrium Ultracentrifugation, and Circular Dichroism studies showed that the extracellular domain of each molecule is a tetramer (Figure 1.4) stabilized by an α-helical neck.

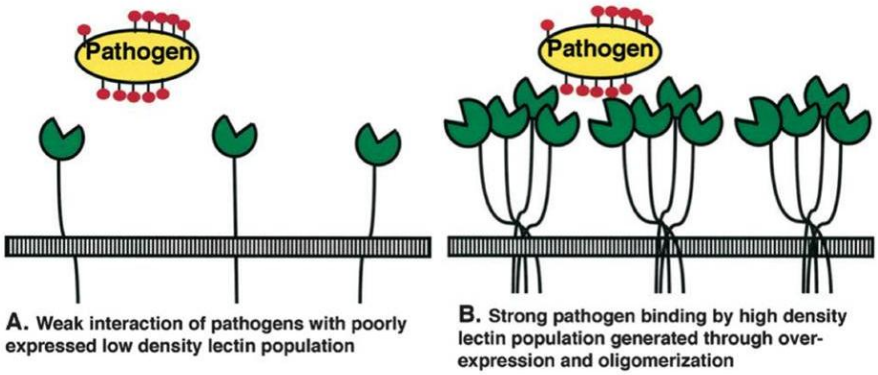


Figure 1.4 *Schematic representation of lectin/pathogen interaction affinity increase due to the avidity phenomenon*.

Such individual CRD has high affinity for mannose-containing oligosaccharides. When the CRDs are clustered in the tetrameric extracellular domain, their arrangement provides a means of amplifying specificity for multiple high mannose oligosaccharides spaced at appropriate distances on the surface.[14] Mannan was the most potent inhibitor (IC50, 6,0 μg/ml), followed by L-fucose (K_i, 6mM); α-methyl D-mannoside (K_i, 15mM), D-mannose (K_i, 23mM); and *N*-acetylglucosamine (K_i, 70mM). Human IgE, sialic acid, and mannose 6-phosphate had no effect on binding: none affects gpl20 binding or gp120-C-type lectin complex binding to CD4.

2.1.2 Key Interaction for viral infection

Most C-type lectins expressed by DC have specificity for mannose-containing carbohydrates. Usually, they recognizes a unique branching and positioning of mannose residues on cell-surface structure, whereas DC-SIGN may recognize high-mannose residues located more internally within a glycan structure and no single terminal mannose residues. This competency offers to DC-SIGN more possibility in the recognition of many pathogens.

The process through which DC-SIGN promotes efficient infection in cells through their CD4/chemokine receptor complex is of particular interest. DC binds strongly to the adhesion molecule ICAM-3 through a veritable "immunological synapse".

ICAM-3, expressed constitutively on the surface of T lymphocytes, carries an abundance of *N*-linked high mannose oligosaccharides just as gp120. Enzymatic removal of *N*-linked carbohydrates from ICAM-3 abrogates its binding to DC-SIGN completely. Moreover, also ICAM-3 binding by DC-SIGN is Ca^{2+}-dependent, main feature of C-type lectins that contain an extracellular CRD. The interaction of DC-SIGN with HIV-1, ICAM-2 and ICAM-3 is inhibited by mannan, mannose, and EGTA.
This observation, deriving above all from many experiments with anti-DC-SIGN antibodies (AZN-D1 and -D2), suggests that ICAM-3 and gp120 bind to similar or closely located, but different sites.

DC-SIGN-ICAM-3 complex, followed by interactions through other adhesion molecules such as LFA-1 and LFA-3 which help in more stable DC–T cell contacts, plays an important role in the activation and proliferation of T cells.[15]

In addition to multiple conformational transitions and interaction of DC-SIGN, crystal structure of a gp120-CD4 complex has revealed that most glycosylation sites within gp120 bind the CD4-binding pocket. So, it is likely that DC-SIGN binding site links to one or more carbohydrate moieties in gp120. It is possible, however, that this C-type lectin domain interacts also with the polypeptide backbone of gp120.

2.1.3 Virus-DC-SIGN Complex Retains Infectivity for a long time

After interaction with the multivalent virus particles, DCs may initiate their activation process. The time course experiment shows that DC-SIGN is able to capture and bind to HIV-1 for more than 4 days, after which the virus can still infect permissive cells. This long-term preservation of HIV-1 in an infectious state would appear to allow sufficient time for it to be transported by DC trafficking from mucosal surfaces to lymphoid compartments, where virus can be transmitted.

During their journey to the regional lymph nodes, DCs protect the viral particles within endocytic vesicles. But other studies are required to understand if this internalization is necessary for infectivity conservation.

2.2 DC-SIGN Or Langerin As Host Receptor

DC-SIGN and Langerin are both C-type lectins of dendritic cells. Equally, they have specificity for mannose and are involved in pathogen recognition. Whereas DC-SIGN allows transinfection of T cells after interaction with viral Env surface, Langerin, on the contrary, contributes to virus elimination. So, it has protective effects on the host cells.

Selective ligands for DC-SIGN, which interact only weakly with Langerin, are potentially useful therapeutic tools against HIV infection. This is the purpose of the latest studies on the mannose- based glycomimetics used as DC-SIGN inhibitor.

2.3 DC-SIGN Or CD4 As Host Receptor

In 2007, studies directed from Jian-Hua Wang and his group[16] analyzed the dendritic cells (DCs) and their key role in the induction and regulation of adaptive immune response. They confirmed DCs are the first cells to encounter HIV at the mucosa, and also discovered that the co-expression of CD4 and DC-SIGN in Raji cells promoted the internalization and intracellular retention of HIV-1. These results suggest that CD4, presents at various levels in DC-SIGN-positive primary cells, is an important regulator of HIV-1 transmission. The principally difference compared to DC-SIGN concerns actually the HIV replication, allowed only from CD4.

Many flow cytometric adhesion assays have demonstrated the ability of HIV-1 coated gp120 to bind to immature DC. The gp120-coated glycans bound efficiently to the DC, and the binding was completely blocked by the anti-DC-SIGN antibodies AZN-D1 and AZN-D2. In contrast, neutralizing anti-CD4 antibodies had no effect on gp120 binding to DC. This result indicates that, although the primary HIV-1 receptor CD4 is expressed on DC, HIV-1 gp120 preferentially binds to DC-SIGN. It is demonstrated the potency of DC in initiating immune responses from resting T cells: small numbers of DCs, and relatively low doses of an antigen or other T cell stimulus, are sufficient to initiate rapid and strong responses, such as T cell proliferation and lymphokine production.

This way, many studies demonstrated that DC-SIGN is a specific dendritic cell surface receptor for the HIV-1 envelope glycoprotein. Finally, DC-SIGN cannot be substituted for CD4 in the process of HIV-1 entry and innate immune response.[17]

2.4 *Monosaccharide Binding Activity*

Synthetic analyses of monosaccharides which interact with CRD of DC-SIGN are represented in the next Table 1[18]:

Table 1

Monosaccharide binding to the extracellular domain fragments of DC-SIGN and DC-SIGNR

Monosaccharide	DC-SIGN		DC-SIGNR	
	K_I	K_I , Sugar/K_I , Man	K_I	K_I , Sugar/K_I , Man
	mM		mM	
Mannose	13.1 ±0.4	1	2.6 ±0.2	1
Methyl α-mannoside	12.5 ±0.5	1.0 ±0.1	3.4 ±0.2	1.0 ±0.1
N-Acetylmannosamine	8.7 ±0.2	0.7 ±0.1	1.3 ±0.1	0.50 ±0.05
Galactose	72 ±5	6.7 ±0.5	20 ±2	8.3 ±0.6
Methyl α-galactoside	270 ±10	27 ±3	112 ±2	34 ±1
Glucose	23 ±1	1.9 ±0.3	7.3 ±0.2	2.6 ±0.1
Methyl α-glucoside	32 ±1	2.6 ±0.3	9.1 ±0.4	2.8 ±0.1
N-Acetylglucosamine	32 ±4	2.5 ±0.3	4.7 ±0.6	1.8 ±0.2
2-Deoxyglucose	28 ±4	2.9 ±0.6	7.3 ±0.3	2.6 ±0.1
L-Fucose	6.7 ±0.5	0.50 ±0.03	4.1 ±0.4	1.7 ±0.1

DC-SIGN binds mannose residue better than they bind galactose, as expected from recent analysis of the CRD sequences. Binding of galactose to mannose-specific C-type CRDs can occur through interactions with the C-1 and C-2 hydroxyl groups of the free sugar. Interaction between mannose residues and DC-SIGN is through equatorial 3-OH and 4-OH groups and the overall binding is maintained by the hydrogen bond and can be mediated or not by water molecules.

In a binding assay of various compounds with rat hepatic lectin and chicken hepatic lectin, the results suggest that the 3- and 4-OH of the target sugar are indispensable, while the 6-OH is not required. Both DC-SIGN and DC-SIGNR show preference for axial C-2 substituents, as they interact with glucose and GlcNAc more weakly than they interact with mannose. For both proteins, affinity was increased by replacement of the axial hydroxyl group of mannose with the *N*-acetyl group in *N*-acetylmannosamine.

2.4.1 High Affinity Binding to High Mannose Oligosaccharides

Recent studies demonstrate that the interaction with high mannose ligands is not limited to terminal mannose residues and rather involves internal residues that stretch along the extended binding site. This aspect was observed in natural ligand interaction with DC-SIGN receptor: it can interact with high mannose-type oligosaccharides but not with single terminal mannose residue.

The main carbohydrate ligand recognized by DC-SIGN is the natural high mannose glycan, $Man_9(GlcNAc)_2$, also represented as Man_9, a branched oligosaccharide which is presented in multiple copies by several pathogen glycoproteins (gp120, GP1, etc.). The $Man_9(GlcNAc)_2$ oligosaccharide competes for binding to the extracellular domain of DC-SIGNR 24-fold more effectively than does mannose. Because the oligosaccharides on natural ligands for DC-SIGN include a range of high mannose structures, binding to these different species was compared by using the extracellular domains of DC-SIGN to probe neoglycolipids coated by high mannose oligosaccharides. Thus, the probe species from $Man_5(GlcNAc)_2$ to $Man_9(GlcNAc)_2$ gave satisfactory results, likely similar to interactions with natural ligand, including the $Man_6(GlcNAc)_2$ found on ICAM-3 isolated from human blood cells.[14a]

Moreover, the increased affinity is often related to the formation of lectin oligomers (as described in paragraph 2.1.1) in which multiple CRDs bind to multiple terminal sugar residues in the ligands, and at the same time oligosaccharides make additional contacts with the beyond the primary mannose-binding site.

2.4.2 Hydrophilic Or Hydrophobic Character Of Ligand

When hydrophobic molecules interact with hydrated binding sites of these lectins, a favorable desolvation and establishment of hydrophobic stacking are often considered as the basis for the hydrophobic interactions. Even though carbohydrates are hydrophilic molecules, they also frequently establish hydrophobic interactions with lectins. Indeed, the two faces (α and β) of the pyranose ring display hydrophobic C–H bonds, which can engage in stacking interactions with protein aromatic side

chains. It is thus common to observe "stacking" of these cycles with aromatic amino acids (Phe, Trp, Tyr), and these interactions are a powerful source of discrimination toward different carbohydrates.

In addition, carbohydrates are generally neutral species, but some sugars (sialic, glucuronic and iduronic acids, aminosugars, alkylated-thiosugars) or modified-sugars (phosphorylated and sulfated sugars) can be either positively or negatively charged. Charged functional groups can give rise to additional contacts with lectins through ionic electrostatic interactions with charged residues.[12b]

Finally, water molecules have a crucial role in carbohydrate–lectin binding events both as solvation molecules and as a relay in the establishment of hydrogen-bonding network. Hydrogen bonding through an intermediate water molecule is indeed quite frequent, and such a relayed hydrogen bond has a comparable energy. The desolvation of water molecules from both binding sites and carbohydrates ligand is very favorable entropically and is commonly invoked as an important driving force for many biomolecular interactions.

The combination of hydrogen bonding, hydrophobic contacts, metal-chelation, and ionic interactions possesses all of the energetic and geometrical features that allow a specific network of interactions as the basis for the reading of the stereochemical information carried by carbohydrates through the "sugar code".

3. THE PATH TOWARDS NEW ANTIRETROVIRAL DRUGS

Since 1980, only 32 antiretroviral drugs were approved for AIDS therapy. They target different stages of the HIV-1 life cycle, such as entry, reverse transcription, integration, and maturation. A complementary approach is now dedicated to the search of inhibitors able to prevent mucosal infection and transmission at primary sites of infection, since the epidemic disease has been largely maintained by mucosal transmission. In parallel, the same suction was considered for other viruses and bacteria, like SIV, Hepatitis C, Ebola, Cytomegalovirus, Dengue, SARS and Mycobacterium tuberculosis interacting with the same HIV target receptor, DC-SIGN.

Although the high avidity generated by clustering low-affinity lectin with monomers into oligomeric structures has made it difficult to design drugs aimed at disrupting protein-carbohydrate interactions, the unusually high affinity between the monomeric DC-SIGN CRDs and high-mannose oligosaccharides suggests that they may be useful targets. So, the latest researches on inhibitory mechanism for the DC-SIGN-virus interactions are described after a brief analysis of preventive and post-infectious therapies already in use.

3.1 Prevention with Vaccines[19]

For many years now, numerous efforts were needed to develop an effective and safe vaccine against AIDS to limit viral transmission. The use of live attenuated viruses is not feasible in humans for safety concerns. Vaccines based on protein subunits are scarcely immunogenic, probably because as the host has glycosylated many cellular proteins, its immune system is unlikely to react to the viral carbohydrate sites, thereby providing a different mechanism to bypass the host's immune protection. So, other approaches are based on DNA vaccines or live recombinant vectors encoding HIV-1 antigens to enhance cell mediated immunity (CMI). For individuals already infected with HIV-1, highly active

antiretroviral therapy (HAART) exists thanks to the extensive development of small molecular weight compounds targeting some key viral processes, such as reverse transcription or protease-dependent Gag-Pol precursor protein cleavage (the major targets of HAART).[20]

An important component of any HIV-1 vaccine approach is however the viral envelope glycoproteins (Env), the major target for neutralizing antibodies (Nabs). The use of monomeric gp120 or peptides derived from the immunodominant V3 loop of gp120 generates type-specific antibodies, but not broadly Nabs. This is likely due to occlusion of conserved V3 epitopes on most circulating isolates. The failure of gp120 to elicit efficiently NAbs has moved design efforts to engineering soluble versions with modified structure of the Env spike to recapitulate properties of the functional trimer. But, clinical trials (as the RV144 trial or the HVTN 505) had never demonstrated efficiently preventative effects, nor a decrease viral load among the vaccinees that became infected by the virus during the trial.

3.2 Other Therapeutic Compounds

Most vaccines and adjuvants are designed to induce immune responses that are stored as immunological memory, to rapidly recognize and clear the pathogen on subsequent exposure. But other numerous therapies against AIDS have been developed.

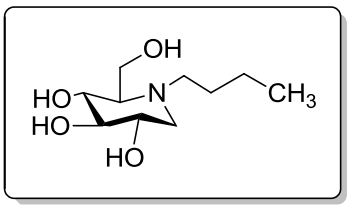


Figure 3.2: *Structure of NB-DNJ.*

For instance, an antiviral currently in use is *N*-butyldeoxynojirimycin (NB-DNJ, represented in Figure 3.2), which has shown to be effective against a range of human viruses including HIV, hepatitis C virus (HCV), hepatitis B virus (HBV), parainfluenza, West Nile and Dengue viruses. Its activity is to inhibit the early trimming of the glucose moieties from $Glc_3Man_9(GlcNAc)_2$ precursors in the endoplasmic reticulum. No effect of NB-DNJ treatment was seen on the kinetics of the interaction between gp120 and CD4. A major mechanism of action of NB-DNJ as an inhibitor of HIV replication is the impairment of viral entry at the level of post-CD4 binding, due to an effect on viral envelope components.

Unfortunately, the 20% of all new HIV-1 infections involve drug-resistant variants; so this could also happen for the NB-DNJ. The combination with nucleoside analogs as dideoxyinosine (DDI), dideoxycytidine (DDC), or azidothymidine (AZT) represents a valid trick and reduce also the yield of reverse transcriptase activity. Moreover, NB-DNJ, tested in combination with common antiretroviral, needs though to excessive concentrations required to achieve significant activity in vivo.[21]

Therefore the initial block of the interaction between DC and virus represents the most advantageous way to escape the viral infection.

3.3 Synthetic Mimetic Ligands of DC-SIGN

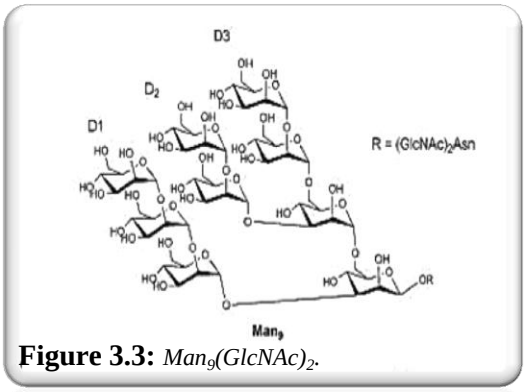


Figure 3.3: *$Man_9(GlcNAc)_2$.*

As already remarked, the affinity of lectins for monovalent carbohydrates is typically weak. DC-SIGN recognizes greatly oligosaccharide structures and has demonstrated a notably monosaccharide affinity for mannose. The affinity for natural oligosaccharide ligand such as $Man_9(GlcNAc)_2$ (showed in Figure 3.3) is a clear example.

Hence, many previous studies reported the synthesis of multivalent carbohydrate-containing molecules and their inhibitor effects. These experiments were involved on different cell lines engineered to express DC-SIGN. The oligosaccharide polymers are specifically characterized by one or many mannose regions, for mimicking the natural ligand Man_9. In this way, the primary carbohydrate-binding site accommodates selectively the equatorial stereochemistry of the C3 and C4 hydroxyls found in the mannopyranosides.[22]

First studies concerned monosaccharide conjugates, linked to a lipid chain of variable length, as represents in Figure 3.4. Their polar mannose epitope moiety binds directly to a hydrophilic linker essential for water solubility; the hydrophilic linker binds to turn hydrophobic saturated lipid chain. It is likely that, in addition to the interaction of the mannose unit with the CRD, the lipid chain is able to interact with some distinct amino acids of DC-SIGN, thus optimizing the binding.[23] Furthermore, it has been determined previously that DC-SIGN exists at the cell surface as a tetramer of identical polypeptide subunits. So in a second time, other findings revealed that multivalent mannose polymers may interfere avidly with the interaction between DC-SIGN and viral Env proteins as affirmed above. The optimized synthesis of many low-molecular-weight compounds gives efficient access to hyperbranched dendritic polymers (one of these represented in Figure 3.4), glyco-dendrimers and glycodendropeptides. The polymerization technique has allowed the binding of more mannose residues (represented in their chair shapes in the Figure 3.5) to a complex branched structure, sometimes characterized by a carrier protein.

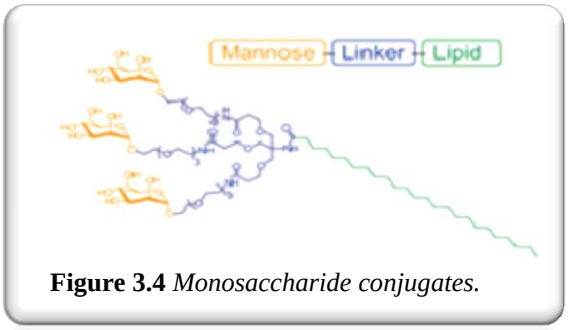
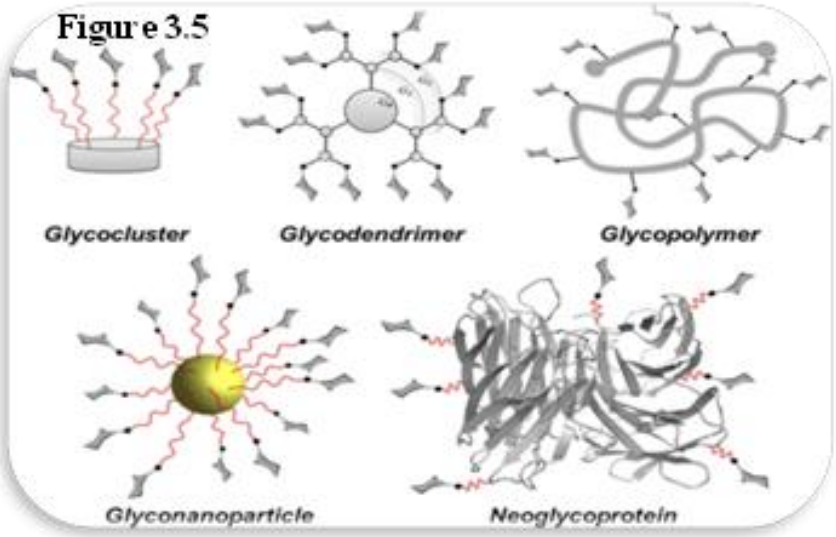
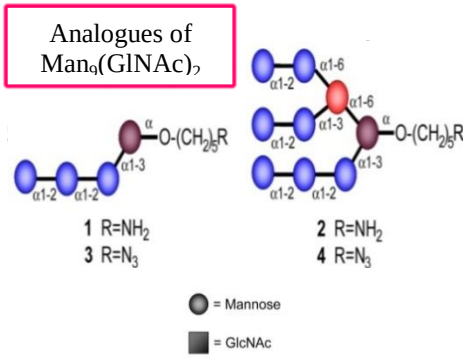
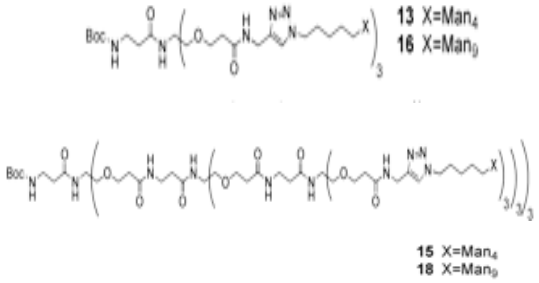
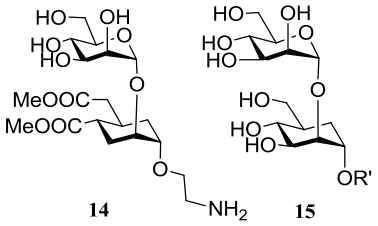
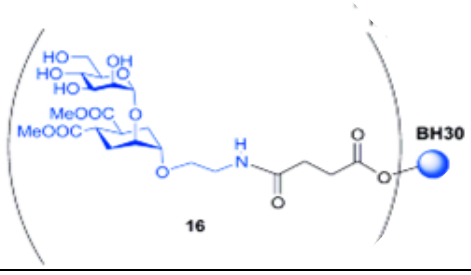
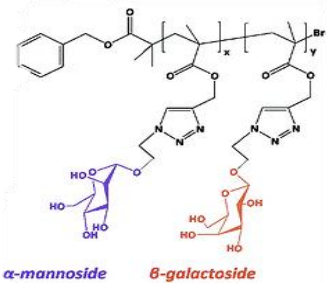


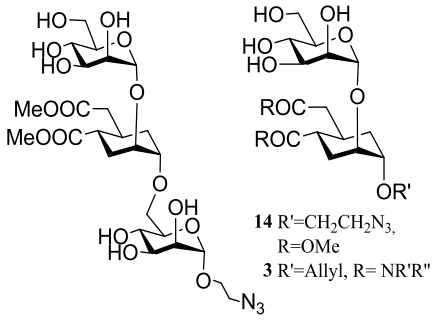
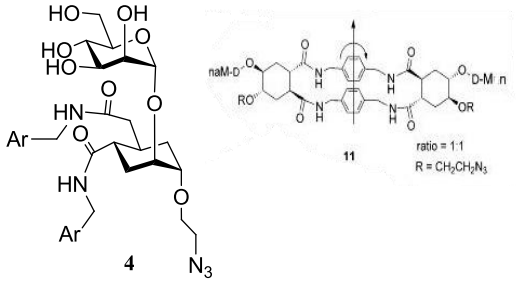
Figure 3.4 *Monosaccharide conjugates.*



It's important to clarify that high affinity of the glycoconjugate molecules requires a saccharide portion (mono, di- or tri-saccharide) that binds tightly to DC-SIGN receptor. On the opposite, the multivalent structures presentation is not necessary, but however important, for obtaining great binding affinity by monovalent compounds. Many studies have, thus, been focused only on activities of low-molecular-weight mannose based glycomimetics. Some glycoconjugates, synthesized by many working groups, are summarized in the following Table 2.

Table 2

YEAR OF PUBBLI-	GLYCOMIMETICS	DERIVATIVES
2007[24]		
2008[25]		
2010[26]		

2013[27]		

3.2.1 Structural Characterization of Pseudo-Mannobiosides

The design of glycomimetic structures is strongly influenced by the natural ligand of DC-SIGN in terms of valence, topology and density of carbohydrates. Man_9 presents in all its arms terminal disaccharides Manα1-2Man which are involved in high mannose recognition process. Thus, many studies demonstrated that unbranched Manα1-2Man-terminated oligosaccharides bind to DC-SIGN almost as efficiently as the entire Man_9. Moreover, the inhibitor activity of modified synthetic analogues was tested in vitro (paragraph 3.2.3).

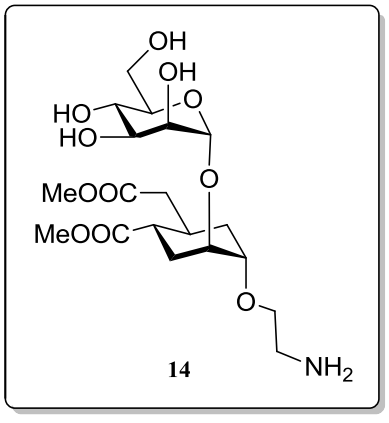
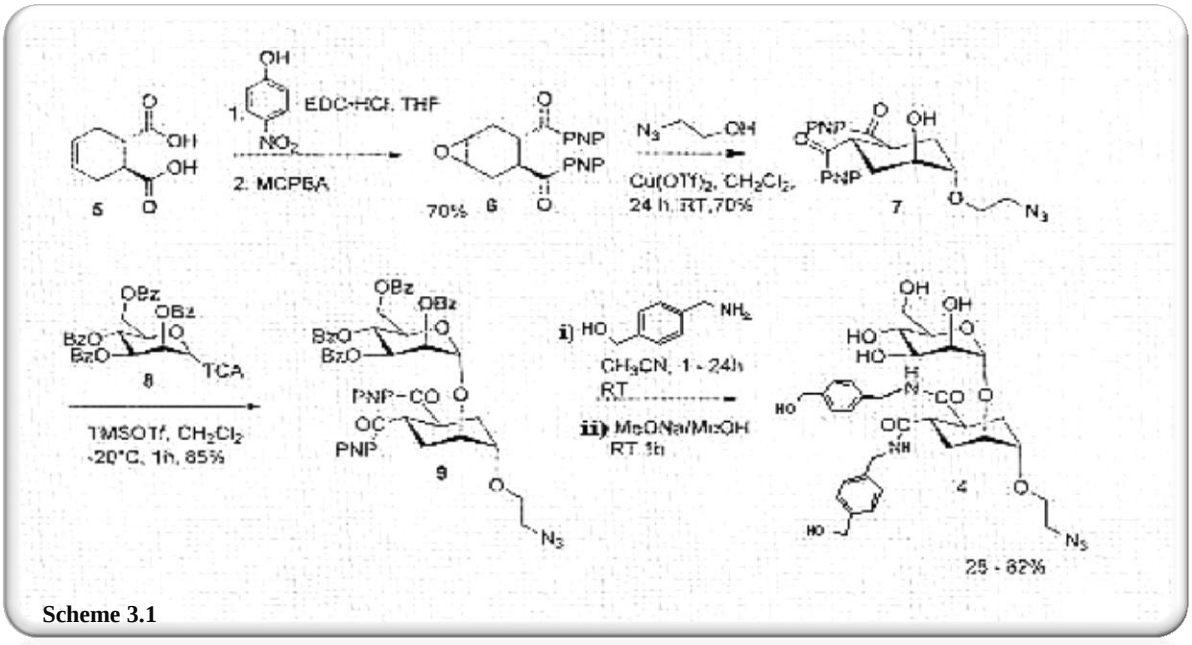


Figure 3.6 *Compound 14, chemical structure of 1'-aminooethyloxi-pseudo-1,2-mannobioside.*

The main features of these 1,2-mannobioside mimics are a mannose unit with free hydroxyl groups, connected to a flexible cyclohexane-based aglycon (Figure 3.6), from which drifts the name of Pseudo-1,2-mannobioside[28] . The cyclohexane portion acts as a mimic of a reducing end mannose residue and features a spacer-arm terminated with amino (Figure 3.6) or azido functionality, useful to generate multivalent DC-SIGN ligands. The resulting pseudo-1,2-mannobioside has the same conformational behavior as the natural disaccharide (Table 2, compound 15), but exhibits also improved stability towards the activity of jack-bean mannosidase.

Recent studies[27], carried out by the same teamwork, show that a strongly DC-SIGN affinity occurs if the R- groups in the position 4' and 5' of the referent compound 2 (Table 2) are lipophilic, for example an aromatic ring. However, most compounds of this recent series possess a low solubility in aqueous media, which limits their applicability. Subsequently, modification of the two ester moieties of the pseudodisaccharide 14 into amides generated glycomimetic ligands with increased activity and a growing water solubility. These molecules provided interesting leads, which could be improved by optimizing the nature of the benzylamide substituents and of appendage in the anomeric position.

First, many types of benzylamine functionalities were replaced during synthetic studies and the relevant benzylamides were tested in adhesion assay. Among them, the best substituent of aromatic portion has proved to be a hydroxymethylene group in *para* position, which appears to play a role as H-bond acceptor. The corresponding methyl ether shows the same inhibitory potency; indeed additional lipophilic groups in the proximity of the acceptor (-OH) or fluorine atoms on the aromatic ring don't improve the affinity. On the other hand, two methoxy substituents in *meta* to the hydroxymethylene group had a negative effect, maybe for a different orientation of the aromatic ring around the *N*-benzylic bond. Their synthesis follows the same procedure of the choosy synthesis of *para*-hydroxymethyl-bis(benzylamide), represented in the following Scheme 3.1.



Scheme 3.1

Starting from enantiomerically pure (1S, 2S)-diacid **5**, the bis(para-nitrophenylester) was synthesized and the double bond oxidized by using MCPBA (*meta*-chloroperbenzoic acid), to afford epoxide **6**. Copper-catalyzed epoxide opening with azidoethanol, followed by mannosylation of alcohol **7** with trichloroacetemidate **8** afforded the pseudodisaccharide **9** in only four steps and with 42 % overall yield. The two activated p-nitrophenyl esters, used as acid-protecting groups throughout the sequence so far, were transformed into amides by reaction with an excess of the appropriate benzylamine (**i)** step) yielding bis(amides) **4** after Zemplén deprotection of the sugar.

3.2.2 Stereochemistry of Cyclohexane Scaffold

In parallel to the previous studies, a diastereoisomer of the compound **4 (**Figure 3.7) with the opposite configuration (1R, 2R) of the cyclohexane ring was synthesized.

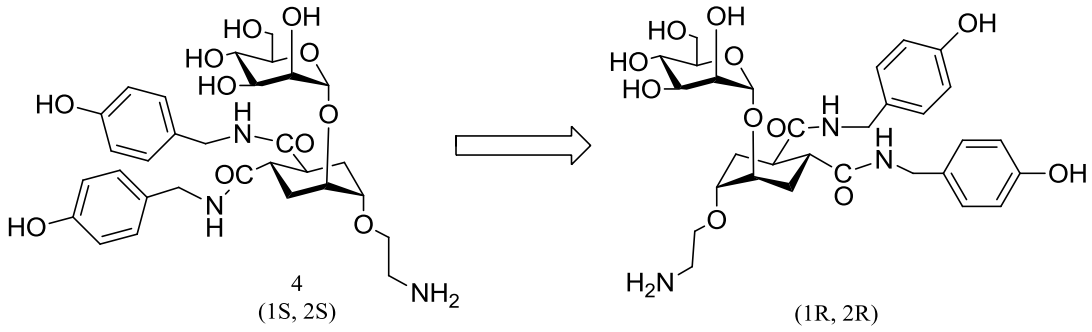


Figure 3.7 *Diastereoisomer of the compound **4** with the opposite configuration (1R, 2R).*

Its synthesis[27b] starts from the (1R,2R)-enantiomer of **5** (Scheme 3.1) and its analysis of ^{1}H NMR spectra confirms that all compounds share the same chair conformation of the cyclohexane ring observed in similar compound **2**. But the relevant pseudodisaccharide didn't demonstrate tightly binding to DC-SIGN, showing that the stereochemistry of the pseudodisaccharide scaffold (like to compound **2**) is important to optimize interaction with the receptor. Only one of analogues synthetized by Professor Anna Bernardi team demonstrates to be the most potent inhibitor (IC_{50} of 31µm),

represented by a divalent macrocyclic ligand (1:1 mixture of isomers), shown in Table **2** (compound 11). Therefore, it is still under consideration due to the low yield with which it was synthetized (18%).

3.2.3 NMR, Docking and Infection studies

Among the compounds mentioned above, the first to be analyzed were the ester derivatives (compound 14 of Table 2). Its interaction with the extracellular domain of DC-SIGN was studied by NMR spettroscopy at 600 MHz. The Saturation Transfer Difference (STD) and the transferred nuclear Overhauser enhancement (TR-NOESY) experiments were also used to observe selectively the binding event of compound 2, substituted with the azido group instead of the amino in the anomeric position, as showed in Figure 3.6. They confirmed that the binding occurred and that the ligand was in close contact with the protein.

A model of the DC-SIGN-compound 2 complex was obtained by docking studies, starting from the PDB structure of the DC-SIGN-Man_4 complex (PDB code IL4).[29] In the QM-polarized ligand docking protocol, ligands are docked with Glide (Grid-based Ligand Docking with Energetics), then charges on the ligand induced by the protein are calculated and a set of the best ligand poses are redocked. This protocol aims to improve the partial charges on the ligand atoms by replacing them with charges derived from quantum mechanical calculations on the ligand in the field of the receptor. Bernardi working group, which synthetized compound 2 and analogues, discovered two rapidly interconverting conformations of the pseudomannobiosides in water solution, the stacked (S) and extended (E) conformations. These two conformations were used as starting points for the docking run. All complexes obtained appeared to maintain the interactions between the Ca^{2+} atom and two hydroxyl groups of the nonreducing end mannose unit.

Finally, infection studies were developed on pseudotype Ebola virus and Jurkat cells, expressing DC-SIGN. The activity of compound 14 in this infection model was compared with the natural disaccharide derivative, presenting the amino-ethoxy linker at the anomeric position (C2'-O-$CH_2CH_2NH_2$). The obtained results showed an IC_{50} of 1.91 mM for the natural disaccharide, value three times higher than the IC_{50} of the mimic **(**0.62 mM). This biological model showed that the designed pseudodisaccharide

14 is a stronger inhibitor than the corresponding disaccharide and taking into account its enzymatic stability could be considered as a promising candidate to prepare multivalent systems to be used as inhibitors of viral infection.[30]

In line with previous findings, similar studies on the benzylic bis(amide) derivatives of general formula 4 (Table 2) were carried out. Since NMR interaction studies of compound 4 showed some overlapping of key signals, a better substrate as its triazole analogue was examined. It proved to be a good lead compound for the analysis. But only its comparison with NMR spectroscopic analysis and STD experiments of methyl ester referent (compound 2) confirmed the same binding mode of two molecules. Really recent structural analysis of the X-ray structure (Figure 3.8, a)) showed that the methyl ester groups of compound 2 in the complex extend towards the DC-SIGN surface in an area in which larger substituents can be accommodated. Larger groups in this position can be expected to reach the protein surface and make additional contacts: binding occurs through coordination of the calcium ion by OH-3 and OH-4 of the mannose residue in an orientation which allows the cyclohexane ring to establish a van der Waals contact with the V_{351} side chain of the receptor; this results in important STD signals for the mannose moiety and for all cyclohexane protons. Moreover, Docking studies (Figure 3.8, b)) suggested that H-bond donors and acceptors out of the plane of the aromatic ring could be beneficial for interaction with the protein surface.

This consideration, together with water solubility and synthetic availability, drove the selection of benzylamines used in synthetic procedures. These molecules were tested for their activity as DC-SIGN ligands by using an SPR (Surface Plasmon Resonance) competition assay. Selectivity against Langerin was also addressed and none of the tested compounds were luckily able to fully inhibit Langerin binding to the surface.

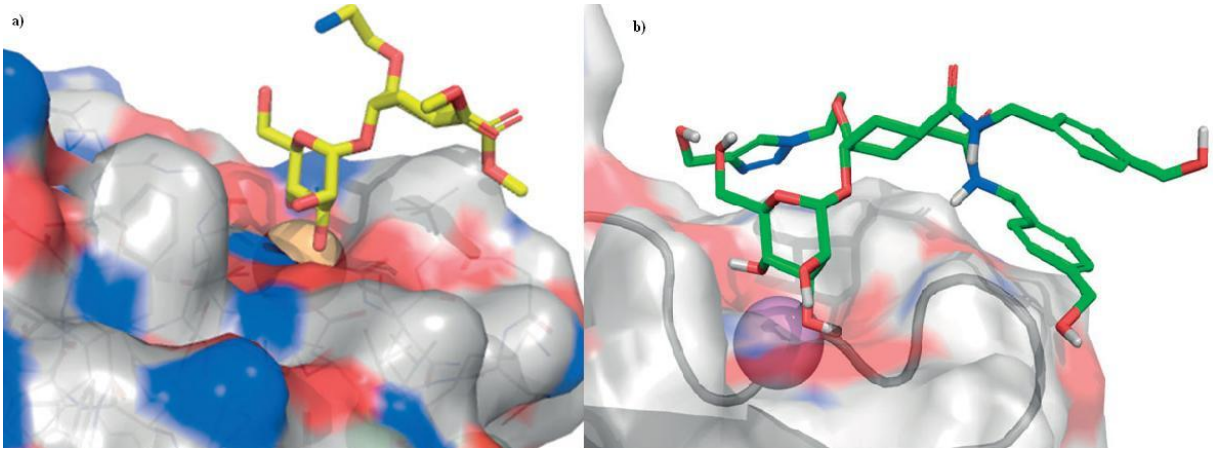


Figure 3.8 *a) X-ray structure of the DC-SIGN/2 complex; b) Docked pose of 13 that qualitatively accounts for the STD data.*

As a result, high saturation in the SDT analysis and great specificity towards DC-SIGN suggested the improved potency of the bis(amides) over the parent dimethyl ester thanks to lipophilic interactions between the aromatic groups of the ligand and the binding site of DC-SIGN.

3.4 Adjuvant of vaccines

Until 2013, DC-SIGN ligands activity was considered only that of inhibition versus target receptor.[31] In an attempt to develop new approaches to vaccine design, many compounds like known receptor ligands or selected antigens could instead exert their effects upon interacting with the innate immune system. Activation of the innate immunity plays a central role in initiating and shaping the adaptive immune response, inducing thus the generation of cytokines, chemokines, and co-stimulatory molecules.

The new adjuvants discovery aims to search for compounds that enhance the vaccine efficiency, even without a clear understanding of how they worked.[32] For instance, analogues of some toll like receptors (TLRs) ligands are already used in licensed vaccines or are in advanced stages of development, providing a starting point for design of such therapeutics.

DC-SIGN activation induces interestingly the stimulation of signal transduction pathways resulting in TLRs signaling, the modulation of immune responses and T cell polarization, with different outcomes depending on the nature of the ligand involved. In the last decade, working group of Milan have nourished much enthusiasm into the field of adjuvant research. They focused on adjuvants interacting with C-type lectin receptors (CLRs), which may be an alternative or complementary to adjuvants targeting TRLs.

Since the function of DC-SIGN is to bind to diverse mannose structures, they have recently demonstrated that a multivalent pseudo-mannosylated compound, Polyman 19 (Figure 3.9) interacts specifically with the carbohydrate-recognition domain of DC-SIGN. It and its glycomimetic derivatives inhibit DC-SIGN mediated HIV-1 infection of cellular and tissue models by competing with the binding of the virus to the receptor.[33]

Polyman 19 (PM 19) is a hexavalent dendrimer carrying six units of a bisamide-based DC-SIGN ligand. The PM 19 efficacy in inhibiting HIV-1 infection was further evaluated exploiting a cervical tissue model. It completely inhibits HIV-1 Bal *trans* infection of $CD4^+$ T cells at 10 μM, and reduces the infection by 50% at 1 μM. PM 19 proved to be more effective than other compounds tested before 2013, such as tetravalent dendron 12, that needed higher concentration to inhibit CD4 T cells *trans* infection.

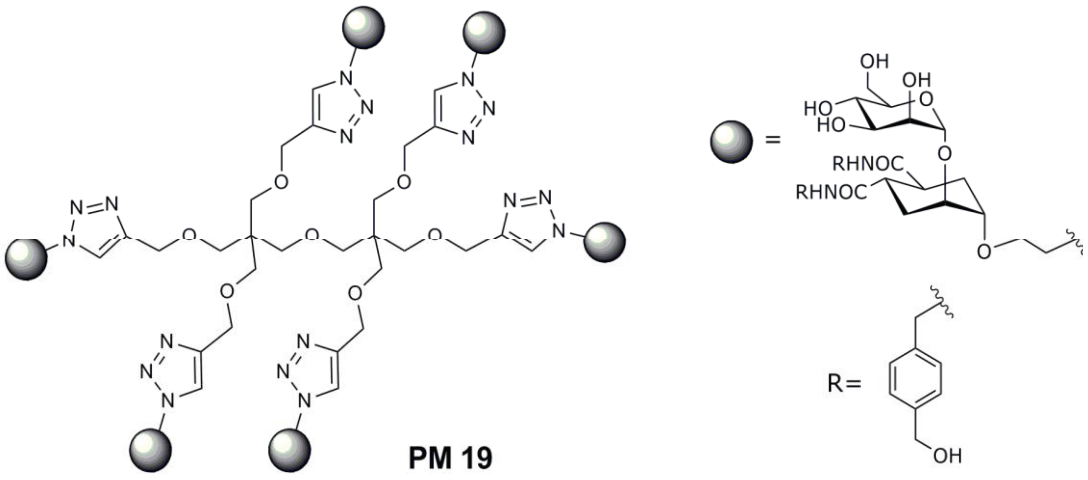


Figure 3.9 *PM19 as adjuvant of vaccines*

In principle, the multivalent structure of these Polyman allows direct conjugation with antigens, realizing a mixed multivalent dendrimers carrying the suitable DC-SIGN ligands and the desired antigen. So this covalent conjugation creates a double effect: DC-SIGN ligands, binding to the receptor, are addressed to endo-lysosomal pathway and then presented on major histocompatibility complex of class II (MHC II) and, in the other hand, the kinked antigen also enhance MHC I presentation through an unknown cross presentation pathway.

The efficacy in inhibiting viral infection by PM 19 and polymannosylated ligands should be further investigated. In addition to the characterization of their role, the purpose of many ongoing studies is to expand the structure knowledge to improve the interactions with the target receptors, DC-SIGN.[25/27a]